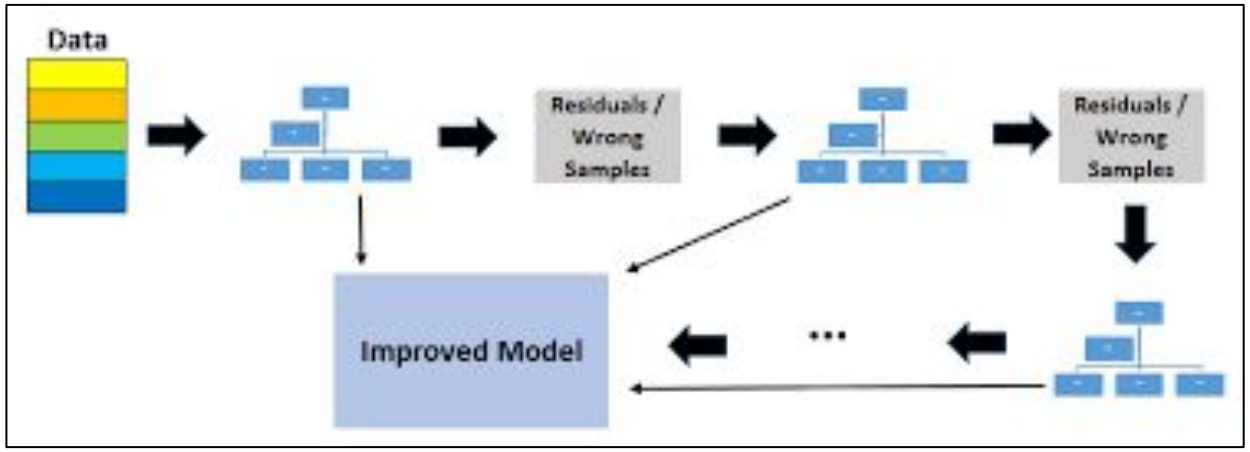
Gradient Boosting Classification

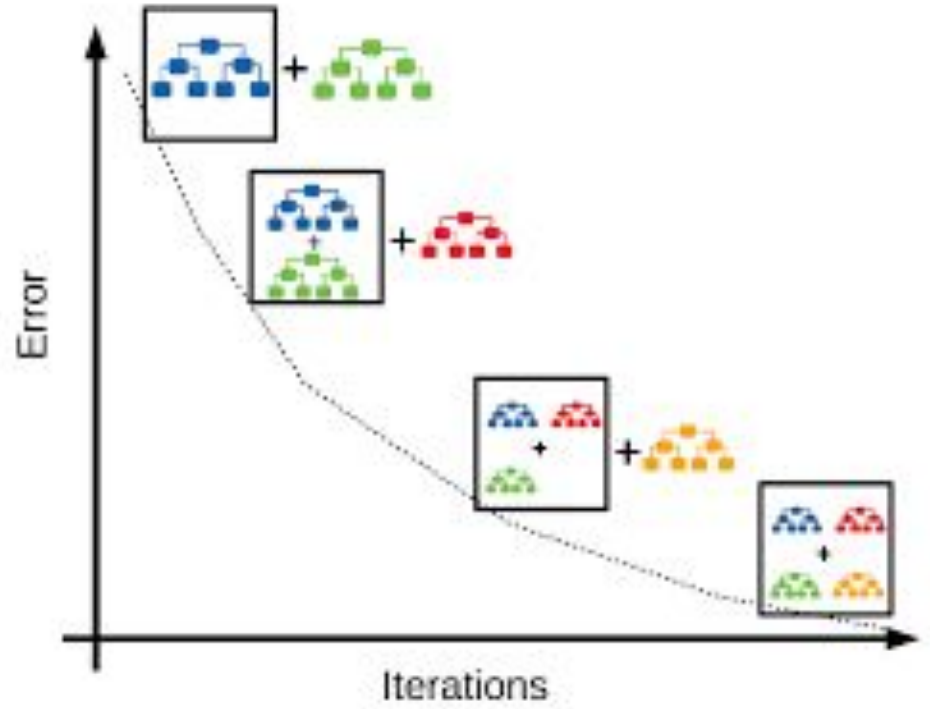
SINDHURA NADENDLA

Ensemble Techniques (More one model)

- Bagging (Random Forest) (Parallelly Multiple trees are created)
- Boosting (GradientBoost) (Corrective Models are created feedback is taken from every model and corrected) (Tree is added Sequentially)

Gradient Boosting Classifier

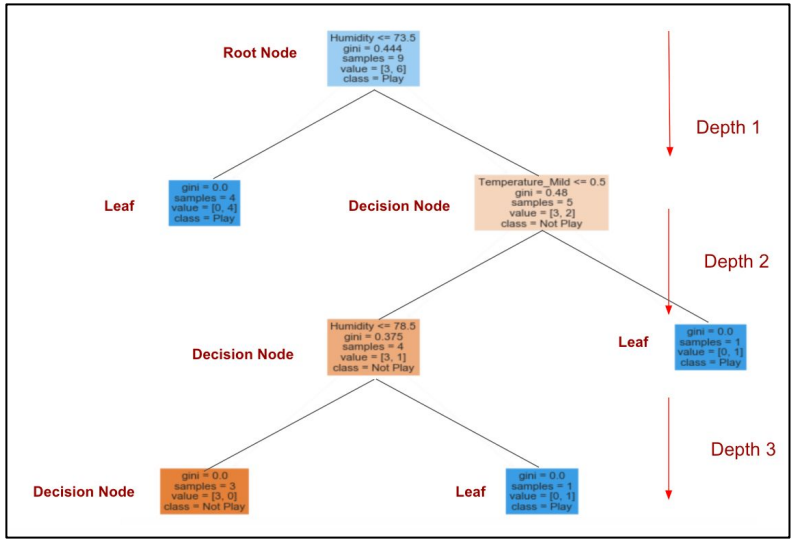




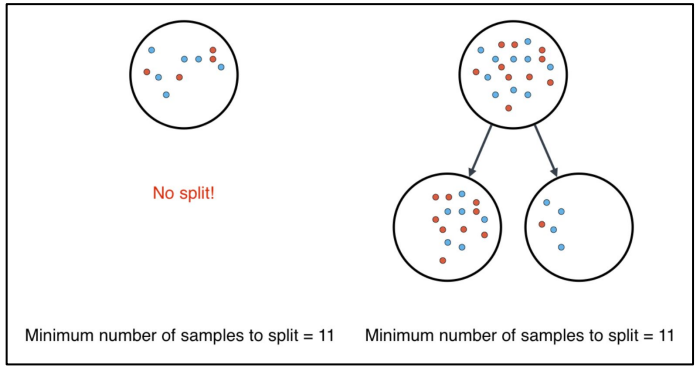
Gradient boosting Parameters

- learning_rate : 0 – 1 (By Default 0.1)
- n_estimators : any (By Default 100)
- max_depth
- min_samples_split
- min_samples_leaf

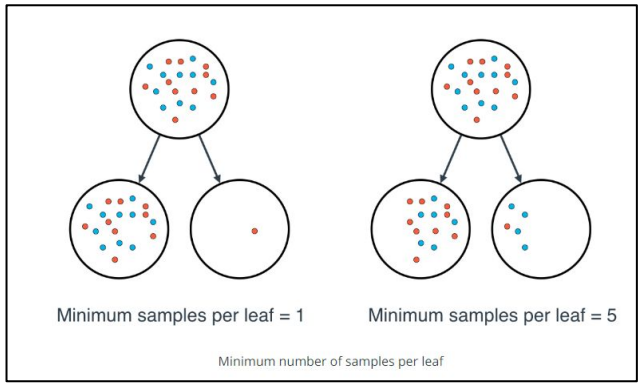
Parameters for Gradient Boost Algo



Max Depth

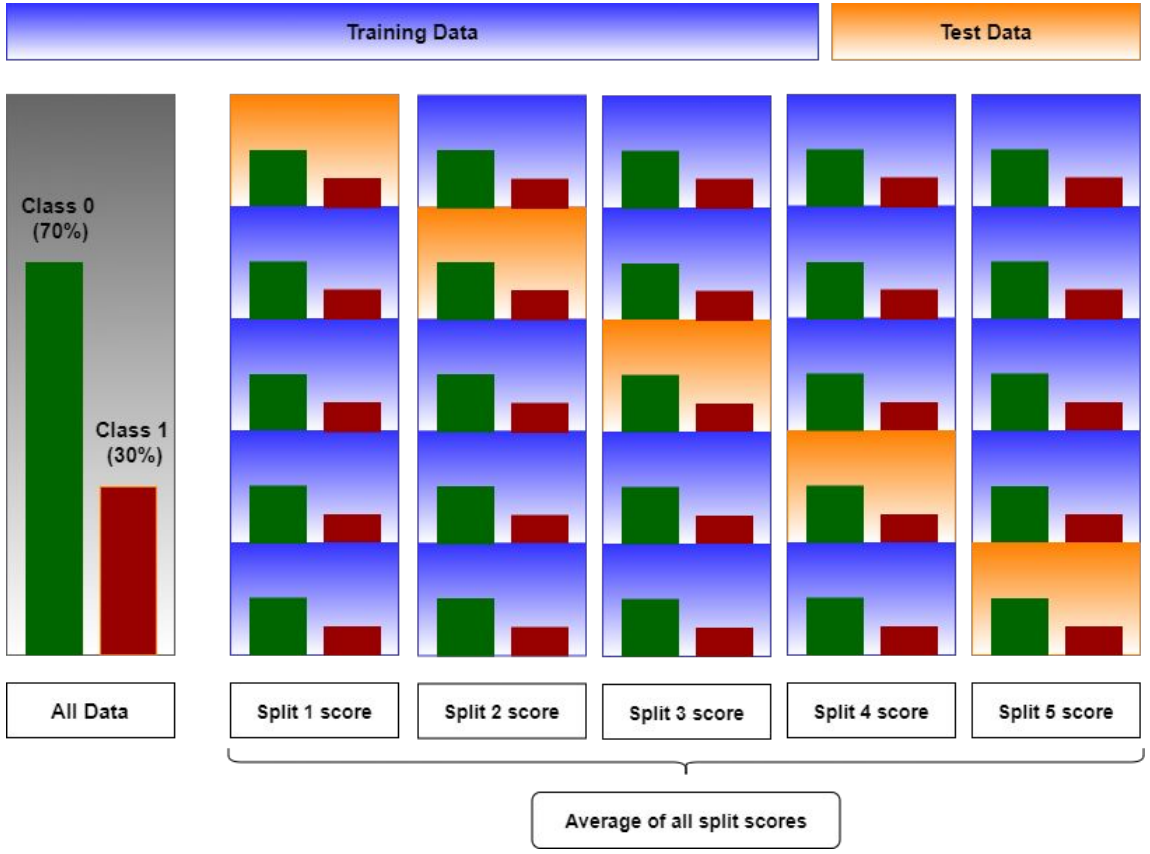


Min Samples Split



Min Samples Leaf

Stratified K-Fold Cross validation



Extreme Gradient Boosting

Advanced , most powerful algorithm which implements Gradient Boosting algorithms, its known for its high accuracy and efficiency in solving both regression and classification problems.

Pros:High accuracy, high performance, prevents overfitting for large datasets , prioritising important features

Cons: Overfitting can occur on smaller datasets due to too many trees, Computationally intensive when training complex models, Sensitive to noise/outliers.

This is an open source provided externally in xgboost library. You need to install it first, to use it.

%pip install xgboost

from xgboost import XGBoostClassifier;

Thank you

UTKARSH GAIKWAD